

# An Evaluation of Robust Cost Functions for RGB Direct Mapping

Alejo Concha and Javier Civera

**Abstract**—The so-called direct SLAM methods have shown an impressive performance in estimating a dense 3D reconstruction from RGB sequences in real-time [1], [2], [3]. They are based on the minimization of an error function composed of one or several terms that account for the photometric consistency of corresponding pixels and the smoothness and the planarity priors on the reconstructed surfaces.

In this paper we evaluate several robust error functions that reduce the influence of large individual contributions to the total error; that most likely correspond to outliers. Our experimental results show that the differences between the robust functions are considerable, the best of them reducing the estimation error up to 25%.

## I. INTRODUCTION

Direct reconstruction or mapping refers to the estimation of a scene 3D structure directly from the *photometric* RGB pixel values of multiple views. This is in opposition to the traditional feature-based techniques that estimate the 3D position of a sparse set of points by minimizing their *geometric* reprojection error. Direct mapping methods have the key advantage of producing denser maps than the traditional feature-based ones, that can only reconstruct salient image points. The minimization of the photometric error of high-gradient pixels produces accurate semi-dense reconstructions [4]. The addition of a regularization term that models scene priors (e.g., smooth surfaces [2], Manhattan or piecewise-planar structure [5]) produces fully dense reconstructions.

The aim of this paper is to explore the role that the error function plays in the accuracy of direct mapping methods. An error function defines how the individual data errors influence the total

error to minimize. Its design has a key importance in the case of spurious data. For example, in the standard  $L^2$  norm, the error grows quadratically. A spurious data point that has a large error also has a large influence in the total cost and can move the estimation apart from the non-spurious data. Error functions with subquadratic growth—even saturated after a certain threshold—can decrease the influence of such outliers.

In our experimental results section we evaluate some of the most common error functions used in the literature. We show that the mean depth error can be reduced up to 25% with an appropriate selection of the robust cost function. Further, according to our experiments, the best error functions are not the ones most commonly used in the literature.

The rest of the paper is organized as follows. Section II describes the related work. Section III details the robust cost functions that we evaluate and section IV details the variational approach to mapping using such robust cost functions. Finally, section V shows the results of our evaluation and section VI presents the conclusions of the paper.

## II. RELATED WORK

The variational approach to the ill-posed problem of optical flow was first proposed in [10]. The algorithm we use in this paper, using the robust  $L^1$  norm for the photometric term, was first proposed in [11] and had the key advantage of resulting in a parallel algorithm suitable for implementation in modern Graphics Processing Units. Early optical flow approaches already noticed the negative effect of outliers and proposed the use of robust cost functions in the data term [12]. Our contribution is the evaluation of such robust cost functions in the 3D mapping problem. Table I references some

Alejo Concha and Javier Civera are with the I3A, Universidad de Zaragoza, {alejocb, jcivera}@unizar.es

	[2]	[5]	[6]	[3]	[1]	[7]	[8], [9]
Photometric	$L^1$	$L^1$	Huber	$L^1$	$L^1$	$L^2$	NCC
Depth Gradient Regularizer	Huber	Huber	Huber	$L^1$	$L^1$	Mean	Huber
Manhattan / Planar Regularizer	-	$L^2$	-	-	-	-	-

TABLE I: Error functions used in the literature.

of the most relevant works on direct RGB mapping and details the error functions they use in the photometric, gradient and Manhattan/piecewise planar regularizers. Notice that the  $L^1$  and Huber norms are the preferred ones. In our evaluation we show how other norms can offer better performance.

[13] evaluated with simulated data the effect that different robust cost functions in the regularization term. We evaluate, using real images, the effect of different cost functions in the data term, the regularization term and the more recent Manhattan/piecewise planar terms (section IV).

### III. ROBUST COST FUNCTIONS

A robust estimate of a parameter vector  $\theta$  is

$$\hat{\theta} = \arg \min_{\theta} \sum_{i=1}^n f(r_i(\theta)) \quad (1)$$

Where  $r_i(\theta) = z_i - g_i(\theta)$  is the  $i^{th}$  residual between the data point  $z_i$  and the data model  $g_i(\theta)$ . Notice that, if  $f$  is defined as the square of the residuals  $f(r_i) = \frac{1}{2}r_i^2$ , the formulation is that of least squares.  $f$  is the error function that should have sub-quadratic growth if we want it to be robust and assign less importance to high-residual points—most likely outliers—than least squares.

The gradient of the objective function  $\sum_{i=1}^n f(r_i(\theta))$  with respect to the parameter vector  $\theta$  is

$$\begin{aligned} \frac{\partial \sum_{i=1}^n f(r_i(\theta))}{\partial \theta} &= \sum_{i=1}^n \frac{\partial f}{\partial r_i}(r_i(\theta)) \frac{\partial r_i}{\partial \theta}(r_i(\theta)) = \\ &= \sum_{i=1}^n \phi(r_i(\theta)) \frac{\partial r_i}{\partial \theta}(r_i(\theta)) \quad (2) \end{aligned}$$

where the derivative  $\phi(r_i(\theta)) = \frac{\partial f}{\partial r_i}(r_i(\theta))$  is usually called the *influence function*. If we define a weight function  $\omega(r_i(\theta))$

$$\omega(r_i(\theta)) = \frac{\phi(r_i(\theta))}{r_i} \quad (3)$$

then equation 2 becomes

$$\frac{\partial \sum_{i=1}^n f(r_i(\theta))}{\partial \theta} = \sum_{i=1}^n \omega(r_i(\theta)) r_i \frac{\partial r_i}{\partial \theta}(r_i(\theta)) \quad (4)$$

Integrating the above equation would give us again the cost function

$$\sum_{i=1}^n f(r_i(\theta)) = \sum_{i=1}^n \int_{\Theta} \omega(r_i(\theta)) r_i \frac{\partial r_i}{\partial \theta}(r_i(\theta)) \partial \theta \quad (5)$$

In order to solve such integral the standard assumption is that the weight is not dependent on the residual and it is assumed constant and taken from the previous iteration ( $k-1$ ) ( $\omega(r_i(\theta)) = \omega(r_i^{(k-1)})$ ).

$$\begin{aligned} \sum_{i=1}^n f(r_i(\theta)) &\approx \sum_{i=1}^n \int_{\Theta} \omega(r_i^{(k-1)}) r_i \frac{\partial r_i}{\partial \theta}(r_i(\theta)) \partial \theta = \\ &= \sum_{i=1}^n \omega(r_i^{(k-1)}) r_i^2(\theta) \quad (6) \end{aligned}$$

With the approximation above, the minimization of equation 1 can be solved as an iteratively reweighted least squares as follows

$$\hat{\theta} = \arg \min_{\theta} \sum_{i=1}^n \omega(r_i^{(k-1)}) r_i^2(\theta) \quad (7)$$

For the complete details on robust statistics the reader is referred to any of the standard books on the topic [14], [15], [16].

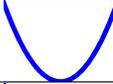
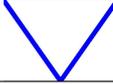
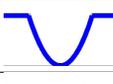
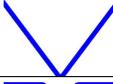
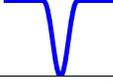
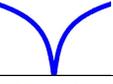
	$f(r)$	$f(r)$ graphical	$\omega(r)$
$L^2$	$\frac{1}{2}r^2$		1
$L^1$	$ r $		$\frac{1}{ r }$
$L^2_{trunc}$	$\begin{cases} \frac{1}{2}r^2 & \text{if }  r  \leq k \\ \frac{k^2}{2} & \text{otherwise} \end{cases}$		$\begin{cases} 1 & \text{if }  r  \leq k \\ 0 & \text{otherwise} \end{cases}$
$L^1_{trunc}$	$\begin{cases}  r  & \text{if }  r  \leq k \\ k & \text{otherwise} \end{cases}$		$\begin{cases} \frac{1}{ r } & \text{if }  r  \leq k \\ 0 & \text{otherwise} \end{cases}$
<i>Huber</i>	$\begin{cases} \frac{r^2}{2} & \text{if } x \leq k \\ k( r - k/2 ) & \text{otherwise} \end{cases}$		$\begin{cases} \frac{1}{ r } & \text{if } x \leq k \\ \frac{k}{ r } & \text{otherwise} \end{cases}$
<i>Tukey</i>	$\begin{cases} \frac{k^2}{6} \left(1 - \left(1 - \left(\frac{r}{k}\right)^2\right)^3\right) & \text{if } x \leq k \\ \frac{k^2}{6} & \text{otherwise} \end{cases}$		$\begin{cases} \left(1 - \left(\frac{r}{k}\right)^2\right)^2 & \text{if } r \leq k \\ 0 & \text{otherwise} \end{cases}$
<i>Cauchy</i>	$\frac{k^2}{2} \log\left(1 + \left(\frac{r}{k}\right)^2\right)$		$\frac{1}{1 + \left(\frac{r}{k}\right)^2}$
<i>Geman-McClure</i>	$\frac{r^2/2}{1+r^2}$		$\frac{1}{(1+r^2)^2}$

TABLE II: Summary of the robust functions evaluated in this paper.

In this work we have evaluated the performance of the most common robust functions in the literature. A summary can be observed in table II. We will take as baselines the norm  $L^2$  –resulting in standard least-squares– and the more robust  $L^1$  and Huber –the most standard ones in state-of-the-art direct mapping.

The truncated  $L^1$  and  $L^2$  norms (in the table as  $L^1_{trunc}$  and  $L^2_{trunc}$ ) are the result of saturating the value of  $f(r)$  for values of  $|r| < k$ . We chose the saturation value  $k = 2\sigma$ , being  $\sigma$  the standard deviation of the error. Due to the presence of gross outliers, we estimated the value of  $\sigma$  robustly from the median value of the distribution  $\sigma = 1.482 \times \text{median}\{\text{median}\{r\} - r_i\}$ .

The Tukey and Geman-McClure functions are very similar to the truncated  $L^2$ , as they behave almost quadratically for small values and saturate for large ones. The Tukey threshold is chosen as  $k = 4.6851$ , the value that achieves 95% rate in the outlier rejection assuming a Gaussian error.

The Huber function is quadratic for small values

and linear for large ones. The limit between the two is  $k = 1.345$ ; again calculated based on a 95% rate spurious rejection for a Gaussian error.

The Cauchy distribution differs from the previous ones in that it has a certain degree of sensitivity to outliers, i.e., the function is not “flat” for large values. The constant is also selected based on a 95% rate spurious rejection ( $k = 2.3849$ ).

#### IV. DIRECT MAPPING

Direct mapping estimates the inverse depth  $\rho(\mathbf{u})$  of every pixel  $\mathbf{u}$  in a reference image  $\mathbf{I}_r$  using the information of such image and several other views  $\mathbf{I}_j$ . The solution comes from the minimization of an energy  $E_\rho$  composed of three terms, a data cost  $\mathbf{C}(\mathbf{u}, \rho(\mathbf{u}))$  based on the photometric error between corresponding pixels, a regularization term  $\mathbf{R}(\mathbf{u}, \rho(\mathbf{u}))$  that models scene priors and a Manhattan term imposing planarity in large untextured areas  $\mathbf{M}(\mathbf{u}, \rho(\mathbf{u}), \rho_p(\mathbf{u}))$ .

$$E_\rho = \int \mathbf{C}(\mathbf{u}, \rho(\mathbf{u})) + \lambda_1 \mathbf{R}(\mathbf{u}, \rho(\mathbf{u})) + \lambda_2 \mathbf{M}(\mathbf{u}, \rho(\mathbf{u}), \rho_p(\mathbf{u})) \quad (8)$$

The photometric error is based on the color difference between the reference image  $\mathbf{I}_r$  and  $m$  other short-baseline views. Every pixel  $\mathbf{u}$  in  $\mathbf{I}_r$  is backprojected at an inverse distance  $\rho$  and projected again in every close image  $\mathbf{I}_j$ .

$$\mathbf{u}^j = \mathbf{T}_{rj}(\mathbf{u}, \rho) = \mathbf{K}\mathbf{R}^\top \left( \left( \begin{array}{c} \mathbf{K}^{-1}\mathbf{u} \\ \frac{\|\mathbf{K}^{-1}\mathbf{u}\|}{\rho} \end{array} \right) - \mathbf{t} \right) \quad (9)$$

The photometric error  $\mathbf{C}(\mathbf{u}, \rho(\mathbf{u}))$  is the summation of the color error  $\epsilon(\mathbf{I}_j, \mathbf{I}_r, \mathbf{u}, \rho)$  between every pixel in the reference image and its corresponding in every other image at an hypothesized inverse distance  $\rho$

$$\mathbf{C}(\mathbf{u}, \rho(\mathbf{u})) = \frac{1}{|I_s|} \sum_{j=1, j \neq r}^m f(\epsilon(\mathbf{I}_j, \mathbf{I}_r, \mathbf{u}, \rho)) \quad (10)$$

$$\epsilon(\mathbf{I}_j, \mathbf{I}_r, \mathbf{u}, \rho) = \mathbf{I}_r(\mathbf{u}) - \mathbf{I}_j(\mathbf{T}_{rj}(\mathbf{u}, \rho)) \quad (11)$$

Notice that we are minimizing a robust function  $f()$  of the error  $\epsilon$ , as defined in section III. We use the robust cost function  $f()$  instead of the weights  $w$  due to the non-convexity of the photometric term—that is minimized by sampling in the literature.

The standard regularizer  $\mathbf{R}(\mathbf{u}, \rho(\mathbf{u}))$  is the Huber norm of the gradient of the inverse depth map  $\|\nabla\rho(\mathbf{u})\|_\epsilon$  and a per-pixel weight  $g(\mathbf{u})$  favoring higher depth gradients for higher color gradients

$$\mathbf{R}(\mathbf{u}, \rho(\mathbf{u})) = g(\mathbf{u})\|\nabla\rho(\mathbf{u})\|_\epsilon \quad (12)$$

We observed that, in this formulation, there is no gain on using robust functions in the regularization. The role of the regularizer is smoothing the large depth gradients that result from a noisy photometric reconstruction. A robust cost function would not reduce the noise. The depth discontinuities that exist in the scene and should be preserved in the estimation are already modeled by the weight  $g(\mathbf{u})$ .

In our experiments we use the Huber norm for the regularizer and compare different alternatives only for the photometric and Manhattan terms.

For man-made scenes, a Manhattan regularizer  $\mathbf{M}(\mathbf{u}, \rho(\mathbf{u}), \rho_p(\mathbf{u}))$  can be added to the gradient-based one  $\mathbf{G}(\mathbf{u}, \rho(\mathbf{u}))$  modeling how far is the estimated inverse depth  $\rho(\mathbf{u})$  from the inverse depth prior  $\rho_p(\mathbf{u})$  coming from the Manhattan or piecewise-planar assumption

$$\mathbf{M}(\mathbf{u}, \rho(\mathbf{u}), \rho_p(\mathbf{u})) = w(\rho(\mathbf{u}) - \rho_p(\mathbf{u}))^2 \quad (13)$$

In this case the error is convex and the gradient of the function is required, therefore we use Iterative Reweighted Least Squares ( $w$ ). The inverse depth prior  $\rho_p(\mathbf{u})$  can be estimated from a region-based multiview reconstruction [17] or from multiview layout estimation in indoor environments [5].

For more details on the energy function  $E_\rho$  and its minimization the reader is referred to [5]. For the initialization of the iterative optimization we will use the photometric depth in the high-gradient image regions and the Manhattan prior for texture-less areas. We have observed that this initialization has better convergence than a photometric one.

## V. EXPERIMENTAL RESULTS<sup>1</sup>

### A. Learning the weighting factors $\lambda_i$

In our model, the relative weights  $\lambda_i$  depend on the inverse depth  $\rho$

$$\lambda_i = \frac{\hat{\lambda}_{i,f}}{1 + 1/\rho} \quad (14)$$

For every robust cost function we learn the optimal value for  $\lambda_{i,f}$  using 5 training sequences. We first re-scale every reconstruction by minimizing the error between the estimated depth  $\frac{1}{\rho(u_j)}$  and the D channel  $D(u_j)$  of every pixel  $u_j$ .

$$\hat{\lambda}_{i,f} = \arg \min_{\lambda_{i,f}} \sum_{k=1}^5 \sum_{u_j=1}^{\#pixels} \left\| \frac{1}{\rho(u_j)} - D(u_j) \right\| \quad (15)$$

<sup>1</sup>A video of the experiments can be watched at <https://youtu.be/PeOux7XhFBI>

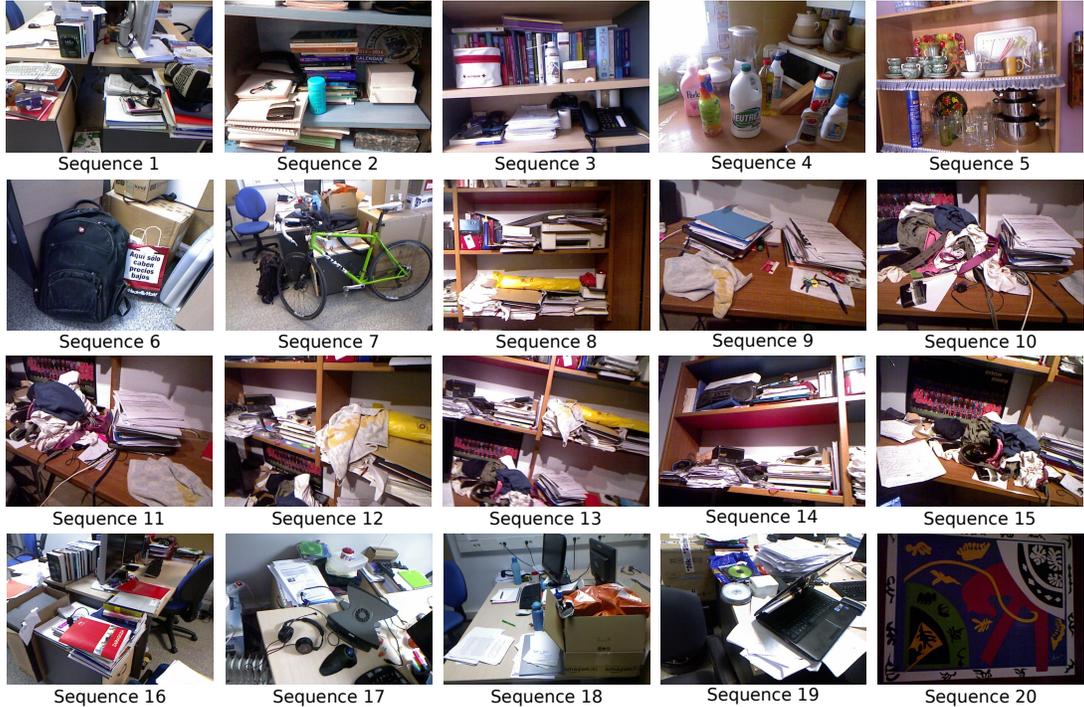


Fig. 1: Sample images for the 20 sequences of our dataset. We selected different indoor scenes including bookshelves (Sequences 2, 3 and 8), bags and backpack (Sequence 6), a bike (Sequence 7), desktops (Sequences 1, 17 and 18) and kitchens (Sequences 4 and 5).

### B. Small scale experiments

We recorded 20 sequences in different indoor environments using a RGB-D sensor. Figure 1 shows the reference frame of each one of the sequences. The depth (D) channel has been used as ground truth and the RGB channel as the input for our algorithm. We chose quite textured scenes in which the standard variational mapping algorithms show a good performance. Direct methods have a low accuracy in low-textured scenes [5] that can hide the influence of the robust error functions.

Table III summarizes the results of our evaluation. Each row corresponds to one experiment –the latest one averages them all– and each column to a different error function.

Look first at the  $L^1$  and  $L^2$  norm results in the 1<sup>st</sup> and 3<sup>rd</sup> column. As expected, the higher weight that the  $L^2$  norm gives to high-residual

points results on a depth map with a larger mean error  $-77$  mm in the latest,  $68$  mm in the former. Notice how the Huber norm (5<sup>th</sup> column), giving linear weight to big errors, behaves similarly to  $L^1$ . The 2<sup>nd</sup> and 4<sup>th</sup> columns show the mean estimation error for the truncated  $L^1$  and  $L^2$  norms. Notice that limiting the maximum cost of large errors has a big influence in the error. Specifically, for the truncated  $L^1$  the error is reduced 25% when compared against the plain  $L^1$ .

The results for the Tukey’s biweight function are displayed in the 6<sup>th</sup> column. The key aspect of this function is that, as the truncated  $L^1$  and  $L^2$ , saturates for high errors. The results are very similar to the truncated  $L^1$ . The same comments apply for the Geman-McClure results in the 8<sup>th</sup> column.

The results for the Cauchy cost function, in the 7<sup>th</sup> column, are between those of the  $L^1$  and

	Mean Depth Error [mm]							
	$L^1$	$L^1_{trunc}$	$L^2$	$L^2_{trunc}$	Huber	Tukey	Cauchy	Geman-McClure
1	82.6	56.2	94.3	57.0	82.6	55.6	68.6	<b>54.4</b>
2	28.6	18.1	37.9	21.9	28.7	<b>17.1</b>	19.7	18.1
3	39.9	31.2	43.2	29.0	39.7	29.1	36.8	<b>27.1</b>
4	36.7	32.2	44.7	33.5	37.0	<b>31.8</b>	33.9	33.8
5	69.5	57.1	77.0	58.1	68.1	<b>56.8</b>	64.7	57.2
6	29.2	<b>25.5</b>	33.0	<b>25.5</b>	28.8	25.7	27.8	27.4
7	91.9	94.5	106.8	91.4	92.3	95.6	<b>89.4</b>	89.7
8	41.5	38.2	45.8	38.6	42.3	37.5	39.1	<b>36.6</b>
9	39.6	40.3	39.4	41.6	39.4	40.3	<b>39.1</b>	41.5
10	45.9	<b>34.2</b>	47.6	35.7	45.1	34.7	43.4	34.8
11	76.3	68.3	78.6	72.1	75.7	66.3	71.4	<b>59.7</b>
12	52.9	49.9	60.3	51.5	52.6	50.0	49.1	<b>48.8</b>
13	55.9	<b>49.6</b>	66.0	50.4	54.6	50.0	55.8	50.6
14	90.1	64.9	65.7	<b>61.3</b>	89.5	69.5	100.4	68.0
15	80.9	46.8	79.8	<b>46.6</b>	79.2	47.4	73.6	48.4
16	110.9	67.9	146.1	72.8	109.6	66.0	86.6	<b>65.2</b>
17	126.3	<b>51.7</b>	160.9	52.3	124.3	54.8	83.9	56.4
18	146.3	<b>95.1</b>	165.3	102.2	145.0	95.2	135.2	103.2
19	120.5	96.4	144.3	106.4	118.4	86.0	101.4	<b>75.6</b>
20	10.3	<b>9.0</b>	9.8	12.9	10.0	<b>9.0</b>	10.5	15.8
Mean	68.8	51.4	77.36	53.2	68.1	50.9	61.5	<b>50.6</b>

TABLE III: Mean errors for several common error functions in the photometric term.

truncated  $L^1$ . The reason is that the Cauchy cost function has a non-zero derivative for large values of the residual, so it still tries to reduce large errors.

The conclusion is that the functions with null derivative for large errors –truncated  $L^1$  and  $L^2$ , Tukey and Geman-McClure– produce more accurate results as they totally ignore large errors and should be preferred over the other ones.

The improvement offered by different functions is not homogeneous. Notice in figure 2 how the depth errors –5<sup>th</sup> and 6<sup>th</sup> columns– are particularly large at depth discontinuities. It is mainly in that regions where the effect of the robust functions is more noticeable. In occlusion-free sequences – e.g., our experiment number 20 imaging a wall– the results are similar for every function.

In figure 2 we show a qualitative comparison for some of the experiments. Notice that the Tukey’s estimation is closer to the ground truth and have a smaller error than the  $L^1$  norm one.

### C. Medium-scale experiment

The aim of this experiment is to evaluate the performance of the robust cost functions in the Manhattan term of our direct mapping approach. Table IV shows the quantitative results. The mean estimation error is reduced around 16% in the first

reference image, 38% in the second one and 13% in the third one with an appropriate error function. As before, those with zero derivative for large errors are the preferred ones.

Figure 3 shows a qualitative view of the results, where the improvement is better explained. 3(b), 3(g) and 3(l) show the layout estimation and labeling –yellow for walls, magenta for clutter and green for floor– for the three reference images considered in this experiment. Notice the large errors of the floor label –some parts corresponding to tables, and even walls in figure 3(l) appear in green color. As a result, the Manhattan energy term (equation 13) is very large there for the right depth. The  $L^2$  norm tries to reduce such high energy, and hence large mapping errors appear. Notice the big differences between the ground truth depth for tables and the  $L^2$  estimated depth in figures 3(d),3(i) and 3(n). Using the Tukey’s function, the large Manhattan errors are assumed spurious and ignored and the rest of the terms define the depth estimation in that area. Notice the more accurate depth estimation in the table areas in figures 3(e),3(j) and 3(o).

## VI. CONCLUSIONS

In this paper we have evaluated several robust cost functions for dense monocular mapping. Our

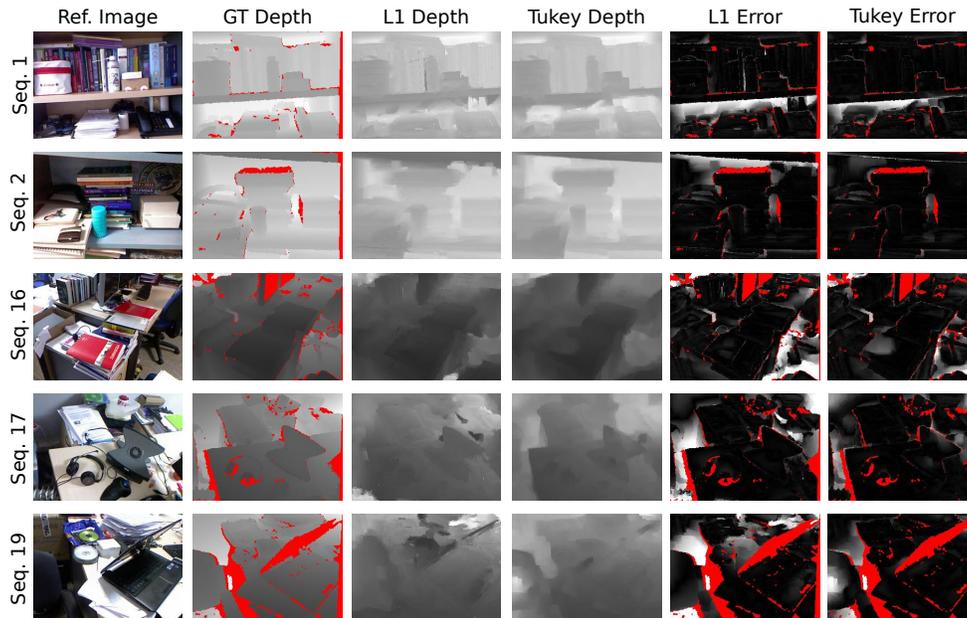


Fig. 2: Five selected quantitative results.  $1^{st}$  column: Reference image.  $2^{nd}$  column: Ground truth depth from a RGB-D sensor. Red stands for no depth data.  $3^{rd}$  column: Estimated depth using the  $L^1$  norm.  $4^{th}$  column: Estimated depth using the Tukey function.  $5^{th}$  column:  $L^1$  depth error, the brighter the larger (and the worse).  $6^{th}$  column: Tukey depth error, the brighter the larger (and the worse).

	Mean Depth Error [cm]							
	$L^1$	$L^1_{trunc}$	$L^2$	$L^2_{trunc}$	Huber	Tukey	Cauchy	Geman-McClure
1	15.69	15.12	16.12	14.30	15.43	<b>13.07</b>	13.73	14.49
2	12.00	11.65	18.52	18.40	11.86	<b>11.06</b>	11.48	11.48
3	32.4	31.59	31.00	<b>29.00</b>	33.4	30.25	29.51	31.91
Mean	20.03	19.45	21.88	20.56	20.23	<b>18.12</b>	18.24	19.29

TABLE IV: Mean errors several common cost functions in the Manhattan term.

results show that the error functions that saturate for large errors –truncated  $L^1$  and  $L^2$ , Tukey and Geman-MacClure– have the best performance. In our experiments the reduction over the standard  $L^1$  and Huber functions is approximately 25% when used in the photometric term and 15% when used in the Manhattan regularization.

#### REFERENCES

- [1] G. Graber, T. Pock, and H. Bischof, “Online 3d reconstruction using convex optimization,” in *2011 IEEE International Conference on Computer Vision Workshops*, 2011, pp. 708–711.
- [2] R. A. Newcombe, S. J. Lovegrove, and A. J. Davison, “DTAM: Dense tracking and mapping in real-time,” in *Computer Vision (ICCV), 2011 IEEE International Conference on*. IEEE, 2011, pp. 2320–2327.
- [3] J. Stühmer, S. Gumhold, and D. Cremers, “Real-time dense geometry from a handheld camera,” in *Pattern Recognition*. Springer, 2010, pp. 11–20.
- [4] J. Engel, T. Schöps, and D. Cremers, “LSD-SLAM: Large-scale direct monocular slam,” in *Computer Vision–ECCV 2014*. Springer, 2014, pp. 834–849.
- [5] A. Concha, W. Hussain, L. Montano, and J. Civera, “Manhattan and piecewise-planar constraints for dense monocular mapping,” in *Robotics: Science and Systems*, July 2014.
- [6] G. Graber, J. Balzer, S. Soatto, and T. Pock, “Efficient minimal-surface regularization of perspective depth maps in variational stereo,” in *CVPR*, 2015.
- [7] J. Engel, J. Sturm, and D. Cremers, “Semi-dense visual odometry for a monocular camera,” in *Computer Vision (ICCV), 2013 IEEE International Conference on*. IEEE,

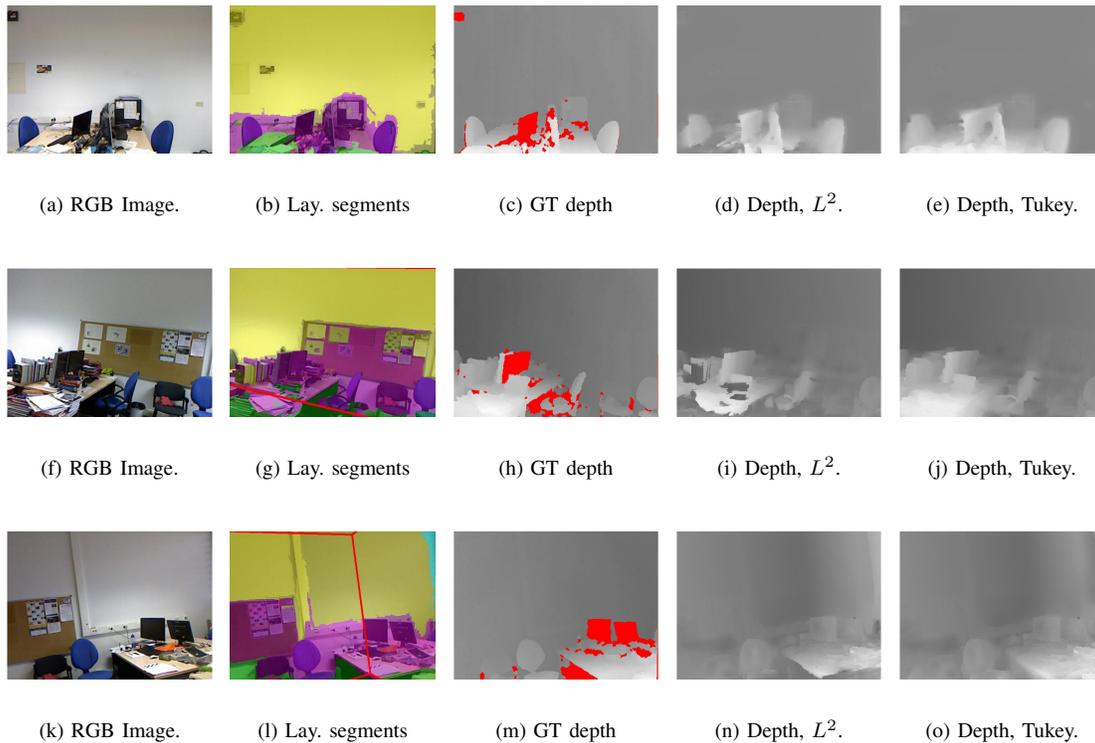


Fig. 3:  $L^2$  and Tukey norms in the Manhattan term. Each row shows a different reference image. 1<sup>st</sup> column: RGB image. 2<sup>nd</sup> column: layout labels (yellow for walls, magenta for clutter and green for floor). Notice that the floor label is wrong. 3<sup>rd</sup> column: ground truth depth. 4<sup>th</sup> column: depth using  $L^2$ . Notice that layout labeling errors produce large depth errors. 5<sup>th</sup> column: depth using Tukey. The large labeling errors have now been weighted out and the depth estimation is closer to the ground truth.

- 2013, pp. 1449–1456.
- [8] M. Pizzoli, C. Forster, and D. Scaramuzza, “Remode: Probabilistic, monocular dense reconstruction in real time,” in *Robotics and Automation (ICRA), 2014 IEEE International Conference on*, 2014, pp. 2609–2616.
- [9] G. Vogiatzis and C. Hernández, “Video-based, real-time multi-view stereo,” *Image and Vision Computing*, vol. 29, no. 7, pp. 434–441, 2011.
- [10] B. K. Horn and B. G. Schunck, “Determining optical flow,” in *1981 Technical Symposium East*. International Society for Optics and Photonics, 1981, pp. 319–331.
- [11] C. Zach, T. Pock, and H. Bischof, “A duality based approach for realtime TV-L 1 optical flow,” in *Pattern Recognition*. Springer, 2007, pp. 214–223.
- [12] M. J. Black and P. Anandan, “A framework for the robust estimation of optical flow,” in *Computer Vision, 1993. Proceedings., Fourth International Conference on*. IEEE, 1993, pp. 231–236.
- [13] L. M. Paz and P. Piniés, “On the impact of regularization for energy optimization methods in dtam,” in *Workshop on Multi View Geometry in Robotics (MVGRO), in conjunction with Robotics: Science and Systems (RSS)*, 2013.
- [14] P. J. Huber, *Robust statistics*. Springer, 2011.
- [15] F. R. Hampel, E. M. Ronchetti, P. J. Rousseeuw, and W. A. Stahel, *Robust statistics: the approach based on influence functions*. Wiley & Sons, 1986.
- [16] R. A. Maronna, R. D. Martin, and V. J. Yohai, *Robust Statistics: Theory and Methods*. Wiley & Sons, 2006.
- [17] A. Concha and J. Civera, “Using superpixels in monocular slam,” in *IEEE International Conference on Robotics and Automation*, Hong Kong, June 2014.